**Архитектуры моделей.**

**Базовая модель 1.**

На вход этой и всех описанных далее моделей подается одноканальное нормализованное изображение логарифмированной мел-спектрограммы речевого сигнала. В базовой модели 1 извлечение признаков производится с помощью сверточной нейронной сети, архитектура которой аналогична архитектуре AlexNet [ссылка], кроме количества входных каналов изображения. Далее извлеченные признаки подаются на блок классификатора, состоящего из 4 полносвязных слоев. В целях регуляризации, после первого полносвязного слоя производится dropout 50% нейронов этого слоя. Схематическое изображение модели представлено на Рисунке 1, описание её слоёв – в Таблице 1.



Рисунок 1. Схема Базовой модели 1.

|  |  |  |
| --- | --- | --- |
| Слой | Параметры | Размерность на выходе |
| Conv1 | Количество фильтров - 64, размер ядра - 11, шаг - 4, пэддинг - 2 | 64х55х55 |
| MaxPool1 | Размер ядра -3, шаг - 2 | 64х27х27 |
| Conv2 | Количество фильтров - 192, размер ядра - 5, шаг - 1, пэддинг - 2 | 192x27x27 |
| MaxPool2 | Размер ядра - 3, шаг - 2 | 192x13x13 |
| Conv3 | Количество фильтров - 384, размер ядра - 3, шаг - 1, пэддинг - 1 | 384x13x13 |
| Conv4 | Количество фильтров - 256, размер ядра - 3, шаг - 1, пэддинг - 1 | 256x13x13 |
| Conv5 | Количество фильтров - 256, размер ядра - 4, шаг - 1, пэддинг - 1 | 256x12x12 |
| MaxPool3 | Размер ядра - 3, шаг - 2 | 256x5x5 |
| FC1 |  | 6400 |
| Dropout | p = 0.5 |  |
| FC2 |  | 2048 |
| FC3 |  | 512 |
| FC4 |  | <количество классов> |

*Таблица 1. Архитектура Базовой модели 1.*

**Базовая модель 2.**

В базовой модели 1 извлечение признаков производится с помощью сверточной нейронной сети, архитектура которой аналогична архитектуре VGG-16 [ссылка], кроме количества входных каналов изображения. Далее извлеченные признаки подаются на блок классификатора, архитектура которого аналогична таковой у Базовой модели 1. Схематическое изображение модели представлено на Рисунке 2, описание её слоёв – в Таблице 2.



Рисунок 2. Схема Базовой модели 2.

|  |  |  |
| --- | --- | --- |
| Слой | Параметры | Размерность на выходе |
| Conv1 | Количество фильтров - 64, размер ядра - 3, шаг - 1, пэддинг - 1 | 64x224x224 |
| MaxPool1 | Размер ядра - 2, шаг - 2, пэддинг - 0 | 64x112x112 |
| Conv2 | Количество фильтров - 128, размер ядра - 3, шаг - 1, пэддинг - 1 | 128x112x112 |
| MaxPool2 | Размер ядра - 2, шаг - 2, пэддинг - 0 | 128x56x56 |
| Conv3 | Количество фильтров - 256, размер ядра - 3, шаг - 1, пэддинг - 1 | 256x56x56 |
| Conv4 | Количество фильтров - 256, размер ядра - 3, шаг - 1, пэддинг - 1 | 256x56x56 |
| MaxPool3 | Размер ядра - 2, шаг - 2, пэддинг - 0 | 256x28x28 |
| Conv5 | Количество фильтров - 512, размер ядра - 3, шаг - 1, пэддинг - 1 | 512x28x28 |
| Conv6 | Количество фильтров - 512, размер ядра - 3, шаг - 1, пэддинг - 1 | 512x28x28 |
| MaxPool4 | Размер ядра - 2, шаг - 2, пэддинг - 0 | 512x14x14 |
| Conv7 | Количество фильтров - 512, размер ядра - 3, шаг - 1, пэддинг - 1 | 512x14x14 |
| Conv8 | Количество фильтров - 512, размер ядра - 3, шаг - 1, пэддинг - 1 | 512x14x14 |
| MaxPool5 | Размер ядра - 2, шаг - 2, пэддинг - 0 | 512x7x7 |
| FC1 |  | 25088 |
| Dropout | p=0.5 |  |
| FC2 |  | 2048 |
| FC3 |  | 512 |
| FC4 |  | <количество классов> |